Transposon-generated 'knock-out' and 'knock-in' gene-targeting constructs for use in mice

Christoph Heiner Westphal and Philip Leder

The conventional technique for targeted mutation of mouse genes entails placing a genomic DNA fragment containing the gene of interest into a vector for fine mapping, followed by cloning of two genomic arms around a selectable neomycin-resistance cassette in a vector containing thymidine kinase [1]; this generally requires 1–2 months of work for each construct. The single 'knock-out' construct is then transfected into mouse embryonic stem (ES) cells, which are subsequently subjected to positive selection (using G418 to select for neomycin-resistance) and negative selection (using FIAU to exclude cells lacking thymidine kinase), allowing the selection of cells which have undergone homologous recombination with the knock-out vector. This approach leads to inactivation of the gene of interest [2]. Recently, an *in vitro* reaction was developed, on the basis of the yeast *Ty* transposon, as a useful technique in shotgun sequencing [3]. An artificial transposable element, integrase enzyme and the target plasmid are incubated together to engender transposition. The DNA is then purified, and subsequently electroporated into bacteria. The transposon and the target plasmid bear distinct antibiotic resistance markers (trimethoprim and ampicillin, respectively), allowing double selection for transposition events. In the present study, we have modified this system to allow the rapid, simultaneous generation of a palette of potential gene targeting constructs. Our approach led from genomic clone to completed construct ready for transfection in a matter of days. The results presented here indicate that this technique should also be applicable to the generation of gene fusion constructs [4–8], simplifying this technically demanding method

Address: Department of Genetics and Howard Hughes Medical Institute, Harvard Medical School, 200 Longwood Avenue, Boston, Massachusetts 02115, USA.

Correspondence: Christoph H. Westphal and Philip Leder.
E-mail: westphal@rascal.med.harvard.edu; leder@rascal.med.harvard.edu

Received: **17 April 1997**
Revised: **19 May 1997**
Accepted: **23 May 1997**

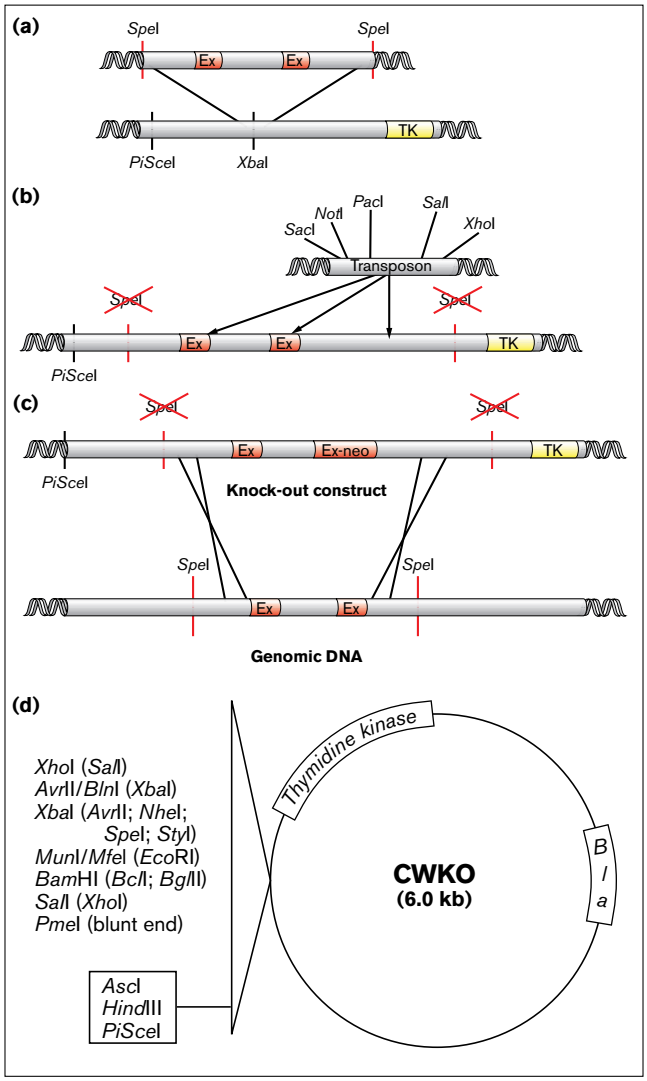
Current Biology 1997, 7:R530–R533
http://biomednet.com/elecref/0960982200700530

Results and discussion

In our transposon-based approach, a genomic fragment containing exon sequences of interest is cloned into a vector containing a number of unique enzyme sites and the *thymidine kinase* gene (TK) at the edge of the multiple cloning site (see Figure 1). As a first step (Figure 1a), enzyme sites at the edge of a genomic clone of interest are destroyed (in the case illustrated here, a genomic *Spe*I fragment is cloned into an *Xba*I site). In the second step (Figure 1b), a simple transposition reaction leads to the random integration of the transposon into the genomic clone. The transposon bears sequences recognized by a number of rare cutters at its extremities, some of which are indicated in Figure 1b. Thousands of unique, individual transposition events can be recovered as distinct doubly resistant colonies from a typical reaction ([3] and our unpublished observations). The desired events — transpositions into the exon of interest — are discerned via a screen that relies on the polymerase chain reaction (PCR) applied to bacterial colonies, using oligonucleotides homologous to the exon DNA. Ligation of a neomycin-resistance (*neo*) cassette into the unique transposon enzyme sites completes the generation of the gene-targeting construct. Neomycin resistance facilitates the selection of homologous recombination events on the basis of regaining external enzyme sites (Figure 1c), and recombinants are verified by Southern blot analysis. Only those constructs undergoing homologous recombination will regain the original enzyme sites at the edge of the construct, at defined distances from the probe. The desired homologous recombinants can then be independently verified by an external probe, if desired. Figure 1d illustrates the vector, named CWKO, that was used for this study, with all the unique sites listed. A genomic fragment containing enzyme sites listed in parentheses will, when cloned into the cognate unique site in this vector, destroy those sites in the genomic DNA. Enzymes that are convenient for linearization of the completed construct are boxed.

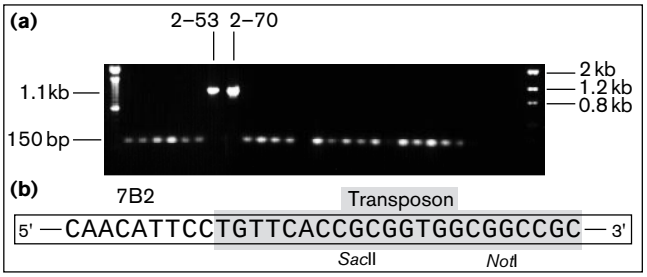
In the present study, we chose to target the mouse neuroendocrine 7B2 gene, the product of which interacts with the prohormone convertase PC2 protein [9]. The 7B2 gene is located within 50 kilobases (kb) of the 3′ end of the *formin* gene (our unpublished observations and [10]). A 7.5 kb genomic *Spe*I fragment was isolated from a bacterial artificial chromosome and cloned into an *Xba*I site in the CWKO vector. A simple transposition reaction, entailing incubation and subsequent phenol extraction steps, was then performed (see Materials and methods section). Colony PCR using oligonucleotides homologous to exon DNA indicated that 4 of 288 transposition events were marked by integration into the exon of interest. Figure 2a shows a representative panel of colony PCR products,

Figure 1



Transposon-mediated generation of mouse 'knock-out' and 'knock-in' vectors. **(a)** Cloning of a representative *Spe*I genomic clone containing two exons (Ex) into an *Xba*I site of the targeting vector, destroying the genomic *Spe*I sites. **(b)** In an *in vitro* transposition reaction, the transposon inserts randomly into the targeting vector. Each ampicillin/trimethoprim-resistant colony represents a unique transposition event. The transposon carries sequences for rare-cutting enzymes at both edges, into which the *neo* cassette is inserted by sticky-end ligation. **(c)** Homologous recombination between the targeting vector and genomic DNA, such that the external *Spe*I sites are regained, allowing the design of a diagnostic genomic Southern blot (see Figure 3 for details). **(d)** A map of the targeting vector CWKO (not drawn to scale). All unique enzyme sites convenient for cloning are indicated. Cognate genomic sites, which can be cloned into the unique vector sites destroying the genomic sites, are indicated in parentheses. Unique enzyme sites useful for targeting vector linearization are boxed. The vector contains the thymidine kinase and ampicillin resistance (*bla*) genes.

Figure 2

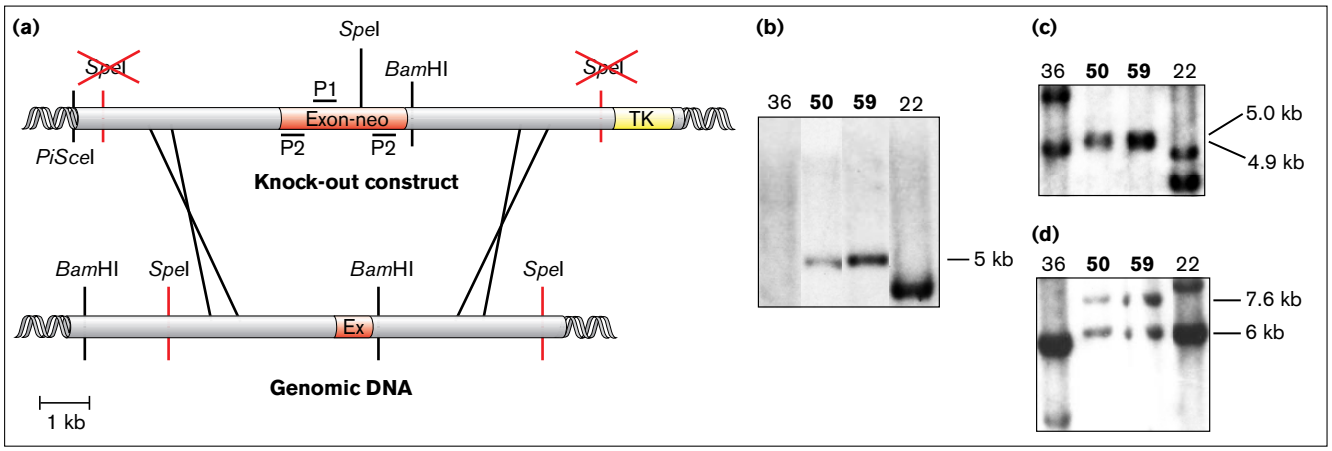


Bacterial colony PCR demonstrates the insertion of the transposon into exon DNA, as verified by sequencing the transposition junction. As described in the text, a genomic *Spe*I clone containing exonic sequences of the mouse neuroendocrine 7B2 gene was cloned into the targeting vector CWKO. An *in vitro* transposition reaction was performed, and resultant ampicillin/trimethoprim-resistant colonies were screened using oligonucleotides capable of amplifying 150 bp of exonic sequences. **(a)** Two of the four clones demonstrated an expected up-shift to 1.1 kb (2–53 and 2–70), indicating that the transposon had inserted within exon 3 of the mouse 7B2 gene. **(b)** The sequence of the transposition site, which was subsequently used to complete the targeting vector. Convenient *Sac*II and *Not*I sites of the transposon, juxtaposed with exon sequences of the 7B2 gene, are labeled.

with the endogenous 150 bp exon band present when integration has taken place elsewhere in the genomic clone. An 'up-shift' to 1.1 kb occurs when the transposon has inserted into exonic DNA, labeled here as 2–53 and 2–70 (note that the transposon is roughly 900 bp). Sequence data, presented in Figure 2b, confirm that transposition has indeed taken place into the 7B2 exon 3. We have confirmed that this approach is generally applicable by generating transpositions into exonic DNA of two other genomic fragments which were used for gene targeting constructs (data not shown). Note that multiple constructs of different genes can be produced simultaneously using this procedure, as a number of transposition reactions may be performed in parallel (Perkin Elmer/ABI routinely perform six or more reactions simultaneously; C. Andre, personal communication).

As a final step, we transfected the knock-out vector generated in the present study into ES cells, as described previously [11]. Linearization of plasmids for transfection into ES cells was achieved with a hypercleavable site for *PiSce*I, which has a 36 bp recognition site [12] and has no reported sites within the mouse genome. We obtained 63 resistant ES clones, of which two were proven to have undergone homologous recombination (Figure 3). Figure 3b shows that ES clones 50 and 59 have the predicted 5 kb band when using a transposon probe (P1 in Figure 3a), whereas ES clone 36 has no detectable band, and ES clone 22 has a band at 4 kb. Figure 3c is a genomic Southern using exon DNA from the mouse 7B2 gene as a probe. This probe is not truly internal, since one is looking to regain two *Spe*I sites, external to the gene-targeting construct, at defined distances. It would be extraordinarily

Figure 3



Homologous recombination is verified by genomic *Spe*I and *Bam*HI digests. **(a)** Homologous recombination leads to regaining of genomic *Spe*I sites (originally destroyed in the targeting vector by cloning into an *Xba*I site; see Figure 1). Genomic *Spe*I and *Bam*HI sites are indicated, as is a *Spe*I site within the *neo* cassette. Transposon DNA (P1) and exonic DNA (P2) were used as probes. As indicated by the sequence shown in Figure 2b, the transposon inserted just after the eightieth base-pair of the 7B2 exon 3. A unique *Not*I site in the transposon was used to clone in the PGK-neo^c bpA cassette. **(b)** The transposon probe identifies a 5 kb band in ES clones 50 and 59, but no band in ES clone 36 and a 4 kb band in ES clone 22. **(c)** The expected 5.0 kb and 4.9 kb doublet on *Spe*I digestion of DNA from clones 50 and 59. All other ES clones showed different banding patterns, represented by that seen in two nontargeted clones, 36 and 22. **(d)** Upon *Bam*HI digestion, clones 50 and 59 show the expected up-shift from the endogenous 6 kb to 7.6 kb, with both bands being of equal intensity. In contrast, all other ES clones showed different banding patterns, and the alternate bands were never of equal intensity (the patterns seen in ES clones 36 and 22 are shown).

unlikely to obtain an insertional event which regained relatively rare-cutting enzyme sites at precisely the same location on both sides of the construct. Since the *neo* cassette contains one *Spe*I site (see Figure 3a), the predicted alteration in the genomic locus is a down-shift from 7.5 kb to a doublet at 5.0 kb and 4.9 kb on genomic *Spe*I digest, which is seen in ES clones 50 and 59 in Figure 3c (note that the *neo* cassette and transposon make the final targeted locus 9.9 kb). This observation is confirmed in Figure 3d, which shows a corresponding up-shift from 6 kb to 7.6 kb upon genomic *Bam*HI digest, using an alternate probe. Note that both knock-out (7.6 kb) and endogenous (6 kb) bands are of equal intensity, indicating that the endogenous locus has been targeted.

The results presented here are based largely on statistical reasoning. We hope that a simple example will illustrate that, given sufficient numbers of random integration events, a number of desired integrants will almost certainly be isolated using this method. Assuming that a given genomic clone is 7.5 kb and contains 375 bp of exonic sequences (genomic DNA is thought to contain roughly 5% exonic sequences), then the chance of one random integration not occurring in the exonic DNA for this example will be the length of all non-exonic vector DNA divided by the length of total DNA, to the first power. Expressed mathematically, this is $((7.5\text{ kb} + 2.5\text{ kb} - 0.375\text{ kb}) / (7.5\text{ kb} + 2.5\text{ kb}))^1$, or $(9{,}625/10{,}000)^1$, since the transposon may also insert in 2.5 kb of the knock-out vector which are not taken up by the ampicillin-resistance cassette. The chance of 100 random integrations not occurring in the exon DNA will, by extension, be $(9{,}625/10{,}000)^{100} = 2\%$. In our laboratory, 300 colony PCR reactions can be set up in under 2 hours, and the likelihood of not recovering a desired insert would then become $(9{,}625/10{,}000)^{300} = 0.001\%$. Note that we and others have routinely obtained thousands of transpositional events per reaction, so that the limiting factor is essentially the number of colony PCR reactions one chooses to perform.

The transposon-mediated approach described here may be generally applicable for the generation of insertional knock-out vectors. The technique is rapid, leading from genomic clone to finished construct in a minimum of four days, and a number of constructs may be generated simultaneously. In addition, different exons in the same genomic clone may be targeted. This can prove useful when different truncations may shed light on the functional significance of distinct protein domains. Finally, the generation of 'knock-in' mice, traditionally an arduous task — in which a portion of the wild-type gene is fused to the cDNA of a heterologous gene — is greatly simplified by the random integration of transposons bearing rare-cutting enzyme sequences. Hence, cloning any cDNA of interest in-frame into a specific genomic locus becomes much less challenging and time-consuming than previously. Several recent papers have indicated the power of such knock-in technology in analyses of the functional

complementation between related genes [4], the use of *lacZ* fusion proteins for precise developmental expression studies [5,13], and the modeling of human cancer translocations [6–8]. We believe that this technique is especially applicable to these types of biological questions, and are currently working on such approaches.

Materials and methods

Construction of the knock-out vector, CWKO

The Invitrogen pSL301 Superlinker plasmid was modified in the following manner. *Hin*dIII and *Not*I sites were filled in using Klenow polymerase. A 36 bp hypercleavable recognition site for *PiSce*I (also containing a *Hin*dIII site) was inserted between *Eco*RI and *Sal*I sites. Oligonucleotide ligation created *Asc*I and *Pme*I sites between the *Sal*I and *Hin*dIII sites. The *thymidine kinase* gene, isolated from the knock-out vector pPNT [1], was blunt-end ligated into a unique *Msc*I site. Diagnostic digestion verified each unique enzyme site listed in Figure 1d.

Generation of transposition events

The manufacturer's protocol (ABI, Perkin Elmer) was followed exactly. Briefly, 200 ng of transposon, 1 μg of target plasmid, integrase, buffer, and water were incubated at 30°C for 1 h. The reaction was stopped with 0.25 M EDTA, 1% SDS, and 5 mg ml^{-1} proteinase K at 65°C for 15 min. After phenol extraction, the product was precipitated with ammonium acetate and isopropanol, washed in 70% ethanol, and resuspended in 10 μl of water; 1 μl was electroporated into highly competent bacterial cells, and plated on selective medium containing 75 μg ml^{-1} ampicillin and trimethoprim. A typical reaction yielded 100–300 colonies per μl, or 1,000–3,000 colonies total. These colonies became apparent on agarose plates 12–15 h after transformation. The only requirements for the transposable element are a 4 bp repeat at either end of a linear, double-stranded DNA molecule [14,15]; random integration into the target DNA engenders only a 5 bp duplication at the site of insertion. Once the transposable element has integrated into the target DNA, it becomes transposition-incompetent.

Colony PCR

Single bacterial colonies were dipped into a master mix containing 0.4 μM primers, 0.2 mM dNTP, 1× PCR buffer, Taq polymerase, and water. The primers used were 5′-AGTTTTCCCAAGAGGACAGG-3′ and 5′-TTCTTCCCACGCTGCAGGG-3′, which amplify exon 3 of the mouse 7b2 gene [16]. Samples were heated to 94°C for 5 min, and then subjected to 30 cycles of 45 sec 94°C, 30 sec 55°C, and 1 min 72°C. After dipping into the PCR master mix, colonies were touched to a master plate, which was incubated at 37°C while PCR and gel analysis was performed. After completion of the PCR reaction in 2.5 h, 1.5% agarose gels were loaded with a multichannel pipettor and run out with markers, to discern the desired transposition events. Setting up 300 PCR reactions, running the PCR program, and loading and analyzing gels was completed in 6–8 h, and colonies from the master plate were then picked and grown up in miniprep format for 8 h. Sticky-end ligation of PGK-neo^c bpA into the targeted exon and subsequent sequencing of the construct was completed in 2 days. Completed constructs were sequenced using a standard protocol (ABI, Perkin Elmer) and analyzed on an ABI 377 automated sequencer.

Homologous recombination in ES cells

ES cell transfection and culture were performed as described previously [11]. Briefly, 40 μg linearized targeting vector was electroporated into ES cells and subjected to positive (G418) and negative (FIAU) selection. Resistant clones were isolated and grown up for DNA, for subsequent analysis by genomic Southern blot.

Acknowledgments

Thanks to Alec Garraway, Mark Bedford, Kevin Fitzgerald and Bill Dietrich for very helpful discussions; to Chaz Andre at ABI for the generous provision of transposon test kits; to Fred Gimble for the kind gift of the *PiSce*I hypercleavable sequence; to Allan Bradley for the PGK-neo^c bpA cassette; to Cathie Daugherty for expert technical assistance with ES cells; and to Heiner Westphal and Rudolf Jaenisch for critical reading of the manuscript. P.L. is a Howard Hughes Medical Institute senior investigator, and C.H.W. is a Sandoz Fellow.

References

1. Tybulewicz VL, Crawford CE, Jackson PK, Bronson RT, Mulligan RC: **Neonatal lethality and lymphopenia in mice with a homozygous disruption of the c-abl proto-oncogene.** *Cell* 1991, **65:**1153-1163.
2. Capecchi MR: **The new mouse genetics: altering the genome by gene targeting.** *Trends Genet* 1989, **5:**70-76.
3. Devine SE, Boeke JD: **Efficient integration of artificial transposons into plasmid targets *in vitro*: a useful tool for DNA mapping, sequencing and genetic analysis.** *Nucleic Acids Res* 1994, **22:**3765-3772.
4. Hanks M, Wurst W, Anson-Cartwright L, Auerbach AB, Joyner AL: **Rescue of the En-1 mutant phenotype by replacement of En-1 with En-2.** *Science* 1995, **269:**679-682.
5. Le Mouellic H, Lallemand Y, Brulet P: **Targeted replacement of the homeobox gene Hox-3.1 by the *Escherichia coli lacZ* in mouse chimeric embryos.** *Proc Natl Acad Sci USA* 1990, **87:**4712-4716.
6. Corral J, Lavenir I, Impey H, Warren AJ, Forster A, Larson TA, *et al.*: **An Mll-AF9 fusion gene made by homologous recombination causes acute leukemia in chimeric mice: a method to create fusion oncogenes.** *Cell* 1996, **85:**853-861.
7. Castilla LH, Wijmenga C, Wang Q, Stacey T, Speck NA, Eckhaus M, *et al.*: **Failure of embryonic hematopoiesis and lethal hemorrhages in mouse embryos heterozygous for a knocked-in leukemia gene CBFB-MYH11.** *Cell* 1996, **87:**687-696.
8. Yergeau DA, Hetherington CJ, Wang Q, Zhang P, Sharpe AH, Binder M, *et al.*: **Embryonic lethality and impairment of haematopoiesis in mice heterozygous for an AML1-ETO fusion gene.** *Nat Genet* 1997, **15:**303-306.
9. Braks JA, Martens GJ: **7B2 is a neuroendocrine chaperone that transiently interacts with prohormone convertase PC2 in the secretory pathway.** *Cell* 1994, **78:**263-273.
10. Wang CC, Chan DC, Leder P: **The mouse formin (Fmn) gene: genomic structure, novel exons, and genetic mapping.** *Genomics* 1997, **39:**303-311.
11. Deng C, Zhang P, Harper JW, Elledge SJ, Leder P: **Mice lacking p21CIP1/WAF1 undergo normal development, but are defective in G1 checkpoint control.** *Cell* 1995, **82:**675-684.
12. Gimble FS, Wang J: **Substrate recognition and induced DNA distortion by the PI-SceI endonuclease, an enzyme generated by protein splicing.** *J Mol Biol* 1996, **263:**163-180.
13. Sosa-Pineda B, Chowdhury K, Torres M, Oliver G, Gruss P: **The Pax4 gene is essential for differentiation of insulin-producing beta cells in the mammalian pancreas.** *Nature* 1997, **386:**399-402.
14. Eichinger DJ, Boeke JD: **The DNA intermediate in yeast *Ty1* element transposition copurifies with virus-like particles: cell-free *Ty1* transposition.** *Cell* 1988, **54:**955-966.
15. Eichinger DJ, Boeke JD: **A specific terminal structure is required for *Ty1* transposition.** *Genes Dev* 1990, **4:**324-330.
16. Braks JA, Broers CA, Danger JM, Martens GJ: **Structural organization of the gene encoding the neuroendocrine chaperone 7B2.** *Eur J Biochem* 1996, **236:**60-67.